



OICE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/047,253

DATE: 02/06/2002  
TIME: 19:09:22

Input Set : A:\CIT1510-4.ST25.txt  
Output Set: N:\CRF3\02062002\J047253.raw

**ENTERED**

2 <110> APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY  
3 COPE, Gregory  
4 VERMA, Rati  
5 ARAVIND, L  
6 KOONIN, Eugene  
7 DESHAIES, Raymond  
9 <120> TITLE OF INVENTION: REGULATION OF TARGET PROTEIN ACTIVITY THROUGH MODIFIER  
PROTEINS

11 <130> FILE REFERENCE: CIT1510-4  
13 <140> CURRENT APPLICATION NUMBER: US/10/047,253  
13 <141> CURRENT FILING DATE: 2002-01-14  
13 <150> PRIOR APPLICATION NUMBER: US 60/261,314  
14 <151> PRIOR FILING DATE: 2001-01-12  
16 <150> PRIOR APPLICATION NUMBER: US 60/322,322  
17 <151> PRIOR FILING DATE: 2001-09-14  
19 <150> PRIOR APPLICATION NUMBER: US 60/322,030  
20 <151> PRIOR FILING DATE: 2001-09-14  
22 <160> NUMBER OF SEQ ID NOS: 22  
24 <170> SOFTWARE: PatentIn version 3.1  
26 <210> SEQ ID NO: 1  
27 <211> LENGTH: 14  
28 <212> TYPE: PRT  
29 <213> ORGANISM: Artificial sequence  
31 <220> FEATURE:  
32 <223> OTHER INFORMATION: JAM domain  
34 <220> FEATURE:  
35 <221> NAME/KEY: MISC\_FEATURE  
36 <222> LOCATION: (1)..(14)  
37 <223> OTHER INFORMATION: Xaa is any amino acid  
39 <400> SEQUENCE: 1  
41 His Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asp  
42 1 5 10  
45 <210> SEQ ID NO: 2  
46 <211> LENGTH: 17  
47 <212> TYPE: PRT  
48 <213> ORGANISM: Artificial sequence  
50 <220> FEATURE:  
51 <223> OTHER INFORMATION: JAM domain  
53 <220> FEATURE:  
54 <221> NAME/KEY: MISC\_FEATURE  
55 <222> LOCATION: (3)..(3)  
56 <223> OTHER INFORMATION: Xaa is Tyr or Ile  
58 <220> FEATURE:  
59 <221> NAME/KEY: MISC\_FEATURE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/047,253

DATE: 02/06/2002

TIME: 19:09:22

Input Set : A:\CIT1510-4.ST25.txt

Output Set: N:\CRF3\02062002\J047253.raw

60 &lt;222&gt; LOCATION: (5)..(5)

61 &lt;223&gt; OTHER INFORMATION: Xaa is Ser or Thr

63 &lt;220&gt; FEATURE:

64 &lt;221&gt; NAME/KEY: MISC\_FEATURE

65 &lt;222&gt; LOCATION: (8)..(16)

66 &lt;223&gt; OTHER INFORMATION: Xaa is any amino acid

68 &lt;400&gt; SEQUENCE: 2

W-OK 70 Gly Trp Xaa His Xaa His Pro Xaa Xaa Xaa Xaa Xaa Ser Xaa Xaa  
 71 1 5 10 15

74 Asp

78 &lt;210&gt; SEQ ID NO: 3

79 &lt;211&gt; LENGTH: 246

80 &lt;212&gt; TYPE: PRT

81 &lt;213&gt; ORGANISM: Homo sapiens

83 &lt;400&gt; SEQUENCE: 3

85 Thr Met Ile Ile Met Asp Ser Phe Ala Leu Pro Val Glu Gly Thr Glu  
 86 1 5 10 15

89 Thr Arg Val Asn Ala Gln Ala Ala Tyr Glu Tyr Met Ala Ala Tyr  
 90 20 25 30

93 Ile Glu Asn Ala Lys Gln Val Gly Arg Leu Glu Asn Ala Ile Gly Trp  
 94 35 40 45

97 Tyr His Ser His Pro Gly Tyr Gly Cys Trp Leu Ser Gly Ile Asp Val  
 98 50 55 60

101 Ser Thr Gln Met Leu Asn Gln Gln Phe Gln Glu Pro Phe Val Ala Val  
 102 65 70 75 80

105 Val Ile Asp Pro Thr Arg Thr Ile Ser Ala Gly Lys Val Asn Leu Gly  
 106 85 90 95

109 Ala Phe Arg Thr Tyr Pro Lys Gly Tyr Lys Pro Pro Asp Glu Gly Pro  
 110 100 105 110

113 Ser Glu Tyr Gln Thr Ile Pro Leu Asn Lys Ile Glu Asp Phe Gly Val  
 114 115 120 125

117 His Cys Lys Gln Tyr Tyr Ala Leu Glu Val Ser Tyr Phe Lys Ser Ser  
 118 130 135 140

121 Leu Asp Arg Lys Leu Leu Glu Leu Leu Trp Asn Lys Tyr Trp Val Asn  
 122 145 150 155 160

125 Thr Leu Ser Ser Ser Ser Leu Leu Thr Asn Ala Asp Tyr Thr Thr Gly  
 126 165 170 175

129 Gln Val Phe Asp Leu Ser Glu Lys Leu Glu Gln Ser Glu Ala Gln Leu  
 130 180 185 190

133 Gly Arg Gly Ser Phe Met Leu Gly Leu Glu Thr His Asp Arg Lys Ser  
 134 195 200 205

137 Glu Asp Lys Leu Ala Lys Ala Thr Arg Asp Ser Cys Lys Thr Thr Ile  
 138 210 215 220

141 Glu Ala Ile His Gly Leu Met Ser Gln Val Ile Lys Asp Lys Leu Phe  
 142 225 230 235 240

145 Asn Gln Ile Asn Ile Ser  
 146 245

149 <210> SEQ ID NO: 4  
 150 <211> LENGTH: 245

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/10/047,253

DATE: 02/06/2002  
 TIME: 19:09:22

Input Set : A:\CIT1510-4.ST25.txt  
 Output Set: N:\CRF3\02062002\J047253.raw

```

151 <212> TYPE: PRT
152 <213> ORGANISM: Homo sapiens
154 <400> SEQUENCE: 4
156 Thr Val Arg Val Ile Asp Val Phe Ala Met Pro Gln Ser Gly Thr Gly
157 1 5 10 15
160 Val Ser Val Glu Ala Val Asp Pro Val Phe Gln Ala Lys Met Leu Asp
161 20 25 30
164 Met Leu Lys Gln Thr Gly Arg Pro Glu Met Val Val Gly Trp Tyr His
165 35 40 45
168 Ser His Pro Gly Phe Gly Cys Trp Leu Ser Gly Val Asp Ile Asn Thr
169 50 55 60
172 Gln Gln Ser Phe Glu Ala Leu Ser Glu Arg Ala Val Ala Val Val Val
173 65 70 75 80
176 Asp Pro Ile Gln Ser Val Lys Gly Lys Val Val Ile Asp Ala Phe Arg
177 85 90 95
180 Leu Ile Asn Ala Asn Met Met Val Leu Gly His Glu Pro Arg Gln Thr
181 100 105 110
184 Thr Ser Asn Leu Gly His Leu Asn Lys Pro Ser Ile Gln Ala Leu Ile
185 115 120 125
188 His Gly Leu Asn Arg His Tyr Tyr Ser Ile Thr Ile Asn Tyr Arg Lys
189 130 135 140
192 Asn Glu Leu Glu Gln Lys Met Leu Leu Asn Leu His Lys Lys Ser Trp
193 145 150 155 160
196 Met Glu Gly Leu Thr Leu Gln Asp Tyr Ser Glu His Cys Lys His Asn
197 165 170 175
200 Glu Ser Val Val Lys Glu Met Leu Glu Leu Ala Lys Asn Tyr Asn Lys
201 180 185 190
204 Ala Val Glu Glu Glu Asp Lys Met Thr Pro Glu Gln Leu Ala Ile Lys
205 195 200 205
208 Asn Val Gly Lys Gln Asp Pro Lys Arg His Leu Glu Glu His Val Asp
209 210 215 220
212 Val Leu Met Thr Ser Asn Ile Val Gln Cys Leu Ala Ala Met Leu Asp
213 225 230 235 240
216 Thr Val Val Phe Lys
217 245
220 <210> SEQ ID NO: 5
221 <211> LENGTH: 421
222 <212> TYPE: PRT
223 <213> ORGANISM: Homo sapiens
225 <400> SEQUENCE: 5
227 Met Pro Asp His Thr Asp Val Ser Leu Ser Pro Glu Glu Arg Val Arg
228 1 5 10 15
231 Ala Leu Ser Lys Leu Gly Cys Asn Ile Thr Ile Ser Glu Asp Ile Thr
232 20 25 30
235 Pro Arg Arg Tyr Phe Arg Ser Gly Val Glu Met Glu Arg Met Ala Ser
236 35 40 45
239 Val Tyr Leu Glu Glu Gly Asn Leu Glu Asn Ala Phe Val Leu Tyr Asn
240 50 55 60
243 Lys Phe Ile Thr Leu Phe Val Glu Lys Leu Pro Asn His Arg Asp Tyr

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/10/047,253

DATE: 02/06/2002  
 TIME: 19:09:22

Input Set : A:\CIT1510-4.ST25.txt  
 Output Set: N:\CRF3\02062002\J047253.raw

```

244 65                               70                               75                               80
247 Gln Gln Cys Ala Val Pro Glu Lys Gln Asp Ile Met Lys Lys Leu Lys
248                               85                               90                               95
251 Glu Ile Ala Phe Pro Arg Thr Asp Glu Leu Lys Asn Asp Leu Leu Lys
252                               100                               105                               110
255 Lys Tyr Asn Val Glu Tyr Gln Glu Tyr Leu Gln Ser Lys Asn Lys Tyr
256                               115                               120                               125
259 Lys Ala Glu Ile Leu Lys Lys Leu Glu His Gln Arg Leu Ile Glu Ala
260                               130                               135                               140
263 Glu Arg Lys Arg Ile Ala Gln Met Arg Gln Gln Gln Leu Glu Ser Glu
264 145                               150                               155                               160
267 Gln Phe Leu Phe Phe Glu Asp Gln Leu Lys Lys Gln Glu Leu Ala Arg
268                               165                               170                               175
271 Gly Gln Met Arg Ser Gln Gln Thr Ser Gly Leu Ser Glu Gln Ile Asp
272                               180                               185                               190
275 Gly Ser Ala Leu Ser Cys Phe Ser Thr His Gln Asn Asn Ser Leu Leu
276                               195                               200                               205
279 Asn Val Phe Ala Asp Gln Pro Asn Lys Ser Asp Ala Thr Asn Tyr Ala
280                               210                               215                               220
283 Ser His Ser Pro Pro Val Asn Arg Ala Leu Thr Pro Ala Ala Thr Leu
284 225                               230                               235                               240
287 Ser Ala Val Gln Asn Leu Val Val Glu Gly Leu Arg Cys Val Val Leu
288                               245                               250                               255
291 Pro Glu Asp Leu Cys His Lys Phe Leu Gln Leu Ala Glu Ser Asn Thr
292                               260                               265                               270
295 Val Arg Gly Ile Glu Thr Cys Gly Ile Leu Cys Gly Lys Leu Thr His
296                               275                               280                               285
299 Asn Glu Phe Thr Ile Thr His Val Ile Val Pro Lys Gln Ser Ala Gly
300                               290                               295                               300
303 Pro Asp Tyr Cys Asp Met Glu Asn Val Glu Glu Leu Phe Asn Val Gln
304 305                               310                               315                               320
307 Asp Gln His Asp Leu Leu Thr Leu Gly Trp Ile His Thr His Pro Thr
308                               325                               330                               335
311 Gln Thr Ala Phe Leu Ser Ser Val Asp Leu His Thr His Cys Ser Tyr
312                               340                               345                               350
315 Gln Leu Met Leu Pro Glu Ala Ile Ala Ile Val Cys Ser Pro Lys His
316                               355                               360                               365
319 Lys Asp Thr Gly Ile Phe Arg Leu Thr Asn Ala Gly Met Leu Glu Val
320                               370                               375                               380
323 Ser Ala Cys Lys Lys Lys Gly Phe His Pro His Thr Lys Glu Pro Arg
324 385                               390                               395                               400
327 Leu Phe Ser Ile Cys Lys His Val Leu Val Lys Asp Ile Lys Ile Ile
328                               405                               410                               415
331 Val Leu Asp Leu Arg
332                               420
335 <210> SEQ ID NO: 6
336 <211> LENGTH: 461
337 <212> TYPE: PRT
338 <213> ORGANISM: Homo sapiens

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/10/047,253

DATE: 02/06/2002  
 TIME: 19:09:22

Input Set : A:\CIT1510-4.ST25.txt  
 Output Set : N:\CRF3\02062002\J047253.raw

```

340 <400> SEQUENCE: 6
342 Met Asp Gln Pro Phe Thr Val Asn Ser Leu Lys Lys Leu Ala Ala Met
343 1 5 10 15
346 Pro Asp His Thr Asp Val Ser Leu Ser Pro Glu Glu Arg Val Arg Ala
347 20 25 30
350 Leu Ser Lys Leu Gly Cys Asn Ile Thr Ile Ser Glu Asp Ile Thr Pro
351 35 40 45
354 Arg Arg Tyr Phe Arg Ser Gly Val Glu Met Glu Arg Met Ala Ser Val
355 50 55 60
358 Tyr Leu Glu Glu Gly Asn Leu Glu Asn Ala Phe Val Leu Tyr Asn Lys
359 65 70 75 80
362 Phe Ile Thr Leu Phe Val Glu Lys Leu Pro Asn His Arg Asp Tyr Gln
363 85 90 95
366 Gln Cys Ala Val Pro Glu Lys Gln Asp Ile Met Lys Lys Leu Lys Glu
367 100 105 110
370 Ile Ala Phe Pro Arg Thr Asp Glu Leu Lys Asn Asp Leu Leu Lys Lys
371 115 120 125
374 Tyr Asn Val Glu Tyr Gln Glu Tyr Leu Gln Ser Lys Asn Lys Tyr Lys
375 130 135 140
378 Ala Glu Ile Leu Lys Lys Leu Glu His Gln Arg Leu Ile Glu Ala Glu
379 145 150 155 160
382 Arg Lys Arg Ile Ala Gln Met Arg Gln Gln Gln Leu Glu Ser Glu Gln
383 165 170 175
386 Phe Leu Phe Phe Glu Asp Gln Leu Lys Lys Gln Glu Leu Ala Arg Gly
387 180 185 190
390 Gln Met Arg Ser Gln Gln Thr Ser Gly Leu Ser Glu Gln Ile Asp Gly
391 195 200 205
394 Ser Ala Leu Ser Cys Phe Ser Thr His Gln Asn Asn Ser Leu Leu Asn
395 210 215 220
398 Val Phe Ala Asp Gln Pro Asn Lys Ser Asp Ala Thr Asn Tyr Ala Ser
399 225 230 235 240
402 His Ser Pro Pro Val Asn Arg Ala Leu Thr Pro Ala Ala Thr Leu Ser
403 245 250 255
406 Ala Val Gln Asn Leu Val Val Glu Gly Leu Arg Cys Val Val Leu Pro
407 260 265 270
410 Glu Asp Leu Cys His Lys Phe Leu Gln Leu Ala Glu Ser Asn Thr Val
411 275 280 285
414 Arg Gly Ile Glu Thr Cys Gly Ile Leu Cys Gly Lys Leu Thr His Asn
415 290 295 300
418 Glu Phe Thr Ile Thr His Val Ile Val Pro Lys Gln Ser Ala Gly Pro
419 305 310 315 320
422 Asp Tyr Cys Asp Met Glu Asn Val Glu Glu Leu Phe Asn Val Gln Asp
423 325 330 335
426 Gln His Asp Leu Leu Thr Leu Gly Trp Ile His Thr His Pro Thr Gln
427 340 345 350
430 Thr Ala Phe Leu Ser Ser Val Asp Leu His Thr His Cys Ser Tyr Gln
431 355 360 365
434 Leu Met Leu Pro Glu Ala Ile Ala Ile Val Cys Ser Pro Lys His Lys
435 370 375 380

```

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/047,253

DATE: 02/06/2002

TIME: 19:09:23

Input Set : A:\CIT1510-4.ST25.txt

Output Set: N:\CRF3\02062002\J047253.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No  
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:41 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:70 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2